Journal of Computer-Aided Molecular Design 2013 vol.27 N8, pages 675-679

Estimation of the size of drug-like chemical space based on GDB-17 data

Polishchuk P., Madzhidov T., Varnek A.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

The goal of this paper is to estimate the number of realistic drug-like molecules which could ever be synthesized. Unlike previous studies based on exhaustive enumeration of molecular graphs or on combinatorial enumeration preselected fragments, we used results of constrained graphs enumeration by Reymond to establish a correlation between the number of generated structures (M) and the number of heavy atoms (N): logM = 0.584 × N × logN + 0.356. The number of atoms limiting drug-like chemical space of molecules which follow Lipinsky's rules (N = 36) has been obtained from the analysis of the PubChem database. This results in M ≈ 1033 which is in between the numbers estimated by Ertl (1023) and by Bohacek (1060). © 2013 Springer Science+Business Media Dordrecht.

http://dx.doi.org/10.1007/s10822-013-9672-4

Keywords

Chemical space, Drug-like chemical space, Graphs enumeration